

Development of a serum biomarker panel predicting recurrence in stage I non–small cell lung cancer patients

Daniel Rinewalt, MD,^a David D. Shersher, MD,^a Shaun Daly, MD,^a Cristina Fhied, MS,^b Sanjib Basu, PhD,^c Brett Mahon, MD,^d Edward Hong, MD,^a Gary Chmielewski, MD,^a Michael J. Liptay, MD,^a and Jeffrey A. Borgia, PhD^{b,d}

Objective: Molecular diagnostics capable of prognosticating disease recurrence in stage I non–small cell lung cancer (NSCLC) patients have implications for improving survival. The objective of the present study was to develop a multianalyte serum algorithm predictive of disease recurrence in stage I NSCLC patients.

Methods: The Luminex immunobead platform was used to evaluate 43 biomarkers against 79 patients with resectable NSCLC, with the following cohorts represented: stage I (T₁–T₂N₀M₀) NSCLC without recurrence (n = 37), stage I (T₁–T₂N₀M₀) NSCLC with recurrence (n = 15), and node-positive (T₁–T₂N₁–N₂M₀) NSCLC (n = 27). Peripheral blood was collected before surgery, with all patients undergoing anatomic resection. Univariate statistical methods (receiver operating characteristics curves and log-rank test) were used to evaluate each biomarker with respect to recurrence and outcome. Multivariate statistical methods were used to develop a prognostic classification panel for disease recurrence.

Results: No relationship was found between recurrence and age, gender, smoking history, or histologic type. Analysis for all stage I patients revealed 28 biomarkers significant for recurrence. Of these, the log-rank test identified 10 biomarkers that were strongly ($P < .01$) prognostic for recurrence. The Random Forest algorithm created a 6-analyte panel for preoperative classification that accurately predicted recurrence in 77% of stage I patients tested, with a sensitivity of 74% and specificity of 79%.

Conclusions: We report the development of a serum biomarker algorithm capable of preoperatively predicting disease recurrence in stage I NSCLC patients. Refinement of this panel might stratify patients for adjuvant therapy or aggressive recurrence monitoring to improve survival. (J Thorac Cardiovasc Surg 2012;144:1344–51)

The American Cancer Society has estimates more than 226,160 new cases of lung cancer, along with approximately 160,340 deaths, in 2012, making it the most common cause of malignancy-related mortality in the United States.¹ Non–small cell lung cancer (NSCLC) is particularly lethal, because many patients present with regional (lymph node) or distant metastases, which has been associated with 24% and 4% 5-year survival, respectively.^{1,2} In contrast, as much as 20% of patients with stage I disease will die of disease recurrence within 5 years of tumor resection.

Recurrence in this group suggests that systemic tumor cell dissemination (locoregional or distant) had already occurred at surgery but was undetected by current clinical and pathologic staging practices.^{3,4} This group of patients with occult “micrometastatic” disease will only receive postoperative surveillance as the standard of care, although these patients might experience substantial clinical benefit from more frequent postoperative surveillance and/or adjuvant systemic chemotherapy (similar to higher stage groups), if appropriate methods were available for definitive selection. The increased use of chest computed tomography (CT) screening protocols in the next several years is expected to shift NSCLC staging demographics toward stage I disease, given the preliminary data from the National Lung Screening Trial, demonstrating a 20% reduction in mortality from NSCLC with screening of high-risk patients relative to chest radiography.⁵ Methods to definitively select stage I patients are urgently needed to help direct patients likely to benefit from adjuvant chemotherapy into prospective clinical trials and to assist with the management of the anticipated increased numbers of stage I NSCLC cases resulting from the implementation of low-dose CT screening protocols.

The objective of the present study was to develop a simple and cost-effective serum test that is prognostic for disease

From the Departments of Thoracic Surgery,^a Biochemistry,^b Preventative Medicine,^c and Pathology,^d Rush University Medical Center, Chicago, Ill.

Supported by awards from the Illinois Chapter of the American Cancer Society (to J.B.), LUNgevity Foundation (to J.B.), Respiratory Health Association of Metropolitan Chicago (to J.B.), Mary Denny Weaver Chair of Cancer Research (to M.L.), and generous philanthropic gifts from the Sapiente Family.

Disclosures: Authors have nothing to disclose with regard to commercial support.

D.R. and D.D.S. had equal contributions to the study.

Read at the 92nd Annual Meeting of The American Association for Thoracic Surgery, San Francisco, California, April 28–May 2, 2012.

Received for publication April 26, 2012; revisions received Aug 1, 2012; accepted for publication Aug 14, 2012; available ahead of print Sept 17, 2012.

Address for reprints: Jeffrey A. Borgia, PhD, Departments of Biochemistry and Pathology, Rush University Medical Center, 554 Cohn Research Bldg, 1735 W. Harrison St, Chicago, IL 60612 (E-mail: Jeffrey_Borgia@Rush.edu).

0022-5223/\$36.00

Copyright © 2012 by The American Association for Thoracic Surgery

<http://dx.doi.org/10.1016/j.jtcvs.2012.08.033>

Abbreviations and Acronyms

CT	= computed tomography
IGFBP	= insulin-like growth factor binding protein
NSCLC	= non-small cell lung cancer
OOB	= out-of-bag
PDIA3	= protein disulfide isomerase 3
PET	= positron emission tomography
TIMP	= tissue inhibitor of metalloproteinase

recurrence in stage I NSCLC patients. For this effort, we reasoned that biomarkers we had reported in our previous studies⁶⁻⁹ as efficacious for identifying patients with locoregional disease progression might also have utility in predicting recurrence in the stage I cohort. We based this hypothesis on the premise that occult micrometastases are the primary cause of disease recurrence in these cases, thereby providing a common mechanistic foundation to investigate. The resulting prognostic classification algorithm represents an exciting advancement in our ability to individualize treatment for stage I NSCLC patients.

PATIENTS AND METHODS

Between 2004 and 2008, we enrolled 79 patients who were divided into the following cohorts: stage I (T_1 - $T_2N_0M_0$) NSCLC without recurrence ($n = 37$), stage I (T_1 - $T_2N_0M_0$) NSCLC with recurrence ($n = 15$), and node-positive (T_1 - T_2N_1 - N_2M_0) NSCLC ($n = 27$). All staging presented was from the pathologic evaluations. The patient inclusion criteria included NSCLC disease confined to the chest without evidence of distant metastases; no preoperative chemo- or radiotherapy within 1 year of our initial blood sampling; and a minimum of 2 years of clinical follow-up data. Patients who died perioperatively were excluded, because they would not have had sufficient follow-up to determine their recurrence status. Demographic information is listed in Table 1. All patient data were acquired with written formal consent and in absolute compliance with the institutional review board at Rush University Medical Center (Chicago, Ill).

Clinical workup at our institution for all NSCLC patients includes CT of the chest and/or abdomen and positron emission tomography (PET). Patients with evidence of possible mediastinal lymph node involvement were evaluated by mediastinoscopy before any additional surgical intervention. All patients with negative findings for mediastinal lymph node involvement on CT-PET scanning or with negative mediastinoscopy findings underwent complete anatomic resection with lymph node dissection of the hilar and/or mediastinal nodes. All intraoperative specimens were examined at surgery by pathologists within our institution to determine margin involvement and potential regional metastatic disease. Routine pathologic examination included hematoxylin and eosin permanent stains on all surgically removed tissue samples with immunohistochemical analysis performed as indicated to confirm disease origin. All samples denoted as lymph node negative had no evidence of metastases or micrometastases on complete pathologic evaluation.

Measurement of Serum Biomarker Concentrations

Serum was prepared using standard phlebotomy protocols from peripheral blood collected in red-top tubes immediately before pulmonary

resection. All specimens were processed within 1 hour of collection, with 50 μ L/mL added to the serum aliquots before archiving at -80°C . No specimen was subjected to more than 2 freeze-thaw cycles before immunoassay evaluation. Assays for insulin-like growth factor binding proteins (IGFBPs) 1 to 7, interferon- γ , CA19-9, cytokeratin 19 fragment 21-1, stem cell factor-1, stromal cell-derived factor-1 α , monocyte colony-stimulating factor, interleukin-2R α , tissue inhibitor of metalloproteinase (TIMP)-1, and IGF-1 were performed using Milliplex Map Kits (EMD-Millipore, Billerica, Mass). The assays used for the circulating autoantibody biomarkers were identical to our previously described methods.⁶⁻⁹ All biomarker concentrations were calculated using a 5-parametric curve fit using xPONENT, version 4.03 (Luminex Corp, Austin, Tex) in a blinded fashion with measurement performed with the FlexMAP 3D system (Luminex Corp). Table 2 lists the 43 biomarkers evaluated in the present study.

The methods for candidate biomarker testing were identical to those previously reported.⁶⁻⁹ In brief, we selected biomarkers for the present study using 2 methods. The first was from a review of the published data for biomarkers that had previously demonstrated efficacy for the early detection of lung cancer.⁶ The second was based on previous efforts in our laboratory.⁸ Specifically, we used Western blots prepared using lysates from a lung adenocarcinoma cell line (HCC827) that were then probed with pooled serum from 2 NSCLC patients groups, the first without nodal involvement and the second with node-positive disease. We compared the immunoblots and performed additional 2-dimensional gels to visualize proteins with a 10-fold difference in immunoreactivity for identification by mass spectrometry. These identified proteins were then selected for analysis in the present study.

Statistical Analysis

Descriptive statistics were obtained, and receiver operating characteristic curves (including area under the curve, specificity, and sensitivity) were used to assess the difference among the 43 individual biomarkers using SPSS statistical software, version 18.0 (SPSS, Chicago, Ill). Fisher's exact and log-rank tests were used to determine significance between the clinical factors (eg, age, gender, race, pack years of smoking exposure, and histologic type) and recurrence. The Kruskal-Wallis test was used to quantify the associations in biomarker concentrations between stage I (with and without recurrence) and positive cohorts, and the Mann-Whitney U and log-rank tests were used to determine the differences related to the interval to recurrence in stage I patients and overall survival. All P values reported are 2-sided. The optimal multivariate panel of biomarkers for predicting the clinical outcome was selected using variable selection algorithms performed within the Random Forest package, as previously described.⁶ In brief, the Random Forest multivariate method grows numerous (1000 in the present study) cross-validated classification trees for a panel of biomarkers, with each tree used to predict group membership for each case. These are counted as the tree "votes" for that group. The forest chooses the group membership having the most votes for all the 1000 trees in the forest. Each such tree is grown by cross-validation, in which a training set (approximately two thirds of the values) is randomly selected from the full data, and each tree is grown on this training data to the largest extent possible (no pruning). The resultant tree is then used to predict the group membership for the remaining validation cases, termed an *out-of-bag* (OOB) prediction. This process is then repeated 1000 times (ie, another training set is randomly selected and a new tree is grown and used to perform another OOB prediction). The classification accuracy of the Random Forest is measured by the averaged error of the OOB predictions across the 1000 trees in the entire forest; this is termed the *OOB error rate*. The OOB error thus uses disjoint subsets of the data for model fitting and validation repeatedly. The reported sensitivity, specificity, and accuracy of the Random Forest method are determined from these OOB predictions from the trees, averaged over the 1000 trees in the forest.

TABLE 1. Demographics for stage I (AJCC 6th edition) patients

Variable	Patients (n)	Recurrence*	Interval to recurrence†
Gender		0.76	0.5
Male	22 (42)		
Female	30 (58)		
Age		0.51‡	0.52§
Range* (y)	50-83		
≤69 y	27 (52)		
>69 y	25 (48)		
Smoking history		1.0	0.85
Smoker	45 (87)		
Nonsmoker	7 (13)		
Race		0.41	0.23
White	42 (81)		
Black	9 (17)		
Hispanic	1 (2)		
Stage (6th edition)			
Ia	18 (35)		
Ib	34 (65)		
pT ₁	18 (35)		
pT ₂	34 (65)		
Histologic subtype		0.28	0.27
Adenocarcinoma	29 (56)		
SCC	16 (31)		
NSCLC, other	7 (13)		
Surgical resection			
Lobectomy	40 (77)		
Pneumonectomy	2 (4)		
Segmentectomy	10 (19)		
(with LN dissection)			

Data in parentheses are percentages. AJCC, American Joint Committee on Cancer; SCC, squamous cell carcinoma; NSCLC, non-small lung cancer; LN, lymph node.

*Fisher's exact test. †Log-rank test. ‡Logistic regression analysis. §Cox proportional hazards regression analysis.

RESULTS

The demographic information for all stage I patients is listed in Table 1. Median patient follow-up for the entire cohort was 40 months, with an average overall survival of 60.5, 35.2, and 28.2 months for the stage I (T₁-T₂N₀M₀) patients without recurrence, with recurrence, and lymph node-positive disease (T₁-T₂N₁-N₂M₀), respectively. After the median 40 months of follow-up, 80% of stage I patients without recurrence were still living compared with only 33% of stage I patients with recurrence and 37% of those with node-positive disease. The median interval to recurrence for the stage I patients was 10.9 months. No significant relationships were discovered between age, gender, race, cancer histologic type, total pack-years smoking history, or PET standard uptake value for any incidence of NSCLC recurrence or interval to recurrence in the stage I cohorts using the Fisher exact test and log-rank test, respectively.

Of the 15 stage I patients with known recurrence, there was a slight female predominance (8/15), with most patients undergoing lobectomy (11/15) or anatomic

segmentectomy with node dissection (4/15). Most of these patients had stage Ib (10/15), with the remaining having stage Ia disease (5/15). No predominant histologic type was noted, with adenocarcinoma (7/15), squamous cell carcinoma (3/15), and not otherwise specified (5/15) represented across this cohort.

Univariate analyses using receiver operating characteristic curve parameters revealed 24 biomarkers with an area under the curve greater than 0.6, with the Mann-Whitney *U* test identifying 3 biomarkers associated with recurrence in stage I patients (Table 2). Additional analyses evaluating biomarker concentrations associated with recurrence in the stage I patient cohort revealed 28 biomarkers significant ($P < .05$) for recurrence using the log-rank test (Table 3). Of these, TIMP-1 and autoantibodies against α -enolase, ubiquitin, survivin, recoverin, peroxiredoxin, methylthioadenosine phosphorylase, protein disulfide isomerase 3 (PDIA3), phosphoglycerate dehydrogenase, and isocitrate dehydrogenase were selected by the log-rank test as being associated with the interval to recurrence ($P < .01$). Moreover, we also identified 21 biomarkers that were significant ($P < .01$) in the stage I cohort in relation to overall survival (Table 3), which included IGFBP-7 and IGFBP-2 and autoantibodies against peroxiredoxin, recoverin, and inosine-5'-monophosphate dehydrogenase. Univariate analysis comparing stage I patients with recurrence and stage I patients with lymph node-positive disease using the Mann-Whitney *U* test revealed 3 biomarkers with differences in expression (receiver operating characteristic > 0.6 ; $P < .05$), including IGFBP-5, IGFBP-7, and autoantibodies against ubiquitin (Table 2).

Multivariate analysis for all stage I patients using the Random Forest algorithm generated a 6-analyte panel consisting of TIMP-1, IGFBP-4, and cytokeratin 19 fragment 21-1, along with autoantibodies against α -enolase, heat shock protein A9, and PDIA3. This Random Forest-generated panel accurately predicted 77% of stage I patients who developed recurrence (with cross-validation). This panel provided 35 cases of true-negative findings, 6 cases of true-positive findings, and 2 cases of false-positive findings in the 52 cases evaluated, for a sensitivity of 74% and specificity of 79%. The area under the curve was calculated as 0.786 (Figure 1).

In the present report, we adhered to the 6th edition of the American Joint Committee on Cancer staging, because this classification scheme reflects the system under which treatment decisions for this cohort of patients was made. Interestingly, 9 patients, all originally classified as having stage Ib with the 6th edition would have been upstaged to stage 2a or 2b using the 7th edition of the staging classification system. Of these patients, 5 developed recurrence, with 3 having been misclassified by our panel. The removal of these patients from the stage I cohort would improve our accuracy to 79%.

TABLE 2. Biomarkers evaluated in all patients

Biomarker	Stage I with recurrence vs LN+		Recurrence in stage I patients	
	<i>P</i> value*	AUC	<i>P</i> value*	AUC
Autoantibodies				
α-Enolase	.168	0.566	.185	0.620
Annexin A1 (ANXA1)	.145	0.557	.284	0.596
Annexin A2 (ANXA2)	.439	0.524	.492	0.438
Calponin 2, isoform b (CNN)	.208	0.556	.332	0.587
c-MYC (MYC)	.106	0.588	.233	0.607
Endoplasmic reticulum protein 29 (ERP-29)	.054	0.579	.106	0.645
Fumarate hydratase (FH)	.723	0.505	.572	0.551
Glyceraldehyde 3 phosphate dehydrogenase (GAPDH)	.203	0.532	.131	0.636
Glyoxalase domain containing 4 (X-3 oxoacid)	.203	0.564	.270	0.600
Heat shock protein A5 (HSPA5)	.423	0.468	.170	0.623
Heat shock protein A9 (HSPA9)	.106	0.520	.031	0.692
Hydroxacyl co-enzyme A dehydrogenase (HADH)	.042	0.566	.094	0.650
Inosine-5'-monophosphate dehydrogenase (IMPDH)	.590	0.537	.777	0.474
Isocitrate dehydrogenase (IDH)	.064	0.562	.082	0.656
Methyl thioadenosine phosphorylase (MTAP)	.168	0.563	.244	0.605
NY-ESO	.101	0.543	.086	0.654
p53	.823	0.438	.171	0.623
Peroxiredoxin (PER)	.121	0.529	.060	0.668
Phosphoglycerate dehydrogenase (PHGDH)	.039	0.558	.036	0.688
Phosphoglycerate mutase 1 (PGAM)	.030	0.595	.146	0.631
Protein disulfide isomerase A3 (PDIA3)	.152	0.500	.039	0.685
Recoverin (REC)	.212	0.499	.094	0.650
Survivin	.254	0.520	.261	0.602
Triose isomerase mutase (TIM)	.138	0.546	.153	0.629
Ubiquilin	.025	0.622	.140	0.632
Circulating factors				
CA19-9	.354	0.547	.847	0.518
Cytokeratin 19 fragment (CYFRA 21-1)	.674	0.555	.461	0.567
Insulin-like growth factor 1 (IGF-1)	.365	0.507	.579	0.550
Insulin-like growth factor-binding protein 1 (IGFBP-1)	.288	0.596	.298	0.594
Insulin-like growth factor-binding protein 2 (IGFBP-2)	.053	0.662	.100	0.648
Insulin-like growth factor-binding protein 3 (IGFBP-3)	.222	0.591	.413	0.574
Insulin-like growth factor-binding protein 4 (IGFBP-4)	.803	0.442	.279	0.597
Insulin-like growth factor-binding protein 5 (IGFBP-5)	.002	0.554	.173	0.623
Insulin-like growth factor-binding protein 6 (IGFBP-6)	.365	0.519	.754	0.529
Insulin-like growth factor-binding protein 7 (IGFBP-7)	.022	0.688	.499	0.561
Interferon (IFN)-γ	.266	0.518	.639	0.458
Interferon-γ-induced protein 10 (IP-10)	.198	0.627	.146	0.631
Interleukin-2 receptor-α (IL-2Rα)	.248	0.579	.525	0.558
Monocyte colony-stimulating factor (M-CSF)	.046	0.608	.553	0.553
Osteopontin	.512	0.553	.606	0.453
Stromal cell-derived factor 1α (SDF-1α)	.030	0.550	.064	0.666
Tissue inhibitor of metalloproteinase 1 (TIMP-1)	.590	0.634	.307	0.593

LN+, Lymph node-positive (disease); AUC, area under the curve. *Mann-Whitney rank sum (2-sided) test.

DISCUSSION

In the present study, we used proven multivariate classification models to develop a serum-based molecular prognostication algorithm that predicts recurrence in stage I NSCLC with greater sensitivity than current clinical and pathologic staging technology. To this end, we formulated a 6-analyte panel capable of predicting recurrence in stage I NSCLC patients. This panel was generated from

peripherally circulating biomarkers with statistically significant concentrations between the 2 stage I cohorts (with and without recurrence). Biomarkers strongly related to the interval to recurrence ($P < .01$) included TIMP-1 and autoantibodies against α-enolase, ubiquilin, survivin, recoverin, peroxiredoxin, methylthioadenosine phosphorylase, PDIA3, phosphoglycerate dehydrogenase, and isocitrate dehydrogenase. Of these, the primary mechanistic

TABLE 3. Median interval to recurrence and overall survival for stage I cohort

Biomarker	Interval to recurrence			Overall survival		
	<i>P</i> value*	Threshold†	Patients (n) ≤ threshold	<i>P</i> value	Threshold†	Patients (n) ≤ threshold
α-Enolase‡	<.001	1156	8	<.046	1156	8
PDIA3‡	.001	202	10	<.038	195	9
PER‡	.002	674	10	<.001	656	9
TIMP-1	.002	419	43	<.020	420	44
PHGDH‡	.004	145	8	<.056	531	40
Ubiquilin‡	.005	383	11	<.110	1749	42
MTAP‡	.006	573	8	<.209	1184	30
IDH‡	.006	306	15	<.123	767	38
Survivin‡	.008	1914	24	<.119	5020	42
REC‡	.010	659	11	<.002	571	10
HSPA9‡	.011	692	28	<.098	753	30
ERP-29‡	.011	144	9	<.039	640	44
X-3 oxoacid‡	<.012	602	12	<.285	1735	41
IP-10	<.014	294	20	<.090	484	38
GAPDH‡	<.014	3007	20	<.022	5701	38
IGFBP-2	<.015	11	16	<.007	14	19
p53‡	<.017	2997	33	<.019	2731	29
SDF-1α	<.017	125	36	<.188	125	36
PGAM‡	<.019	528	40	<.077	530	41
HADH‡	<.023	653	41	<.026	768	43
ANXA1‡	<.024	552	41	<.102	511	39
IGFBP-7	<.025	42	9	<.003	60	43
IMPDH‡	<.031	532	8	<.008	1215	23
MYC‡	<.031	720	10	<.114	1210	17
IL-2Rα	<.031	98	42	<.113	45	25
IGFBP-5	<.033	301	31	<.016	301	31
NYESO‡	<.038	545	9	<.016	523	8
HSPA5‡	<.043	962	27	<.050	1235	31
CNN‡	<.073	442	19	<.026	1607	42
CYFRA 21-1	<.077	999	41	<.101	649	32
TIM‡	<.091	768	11	<.015	590	9
FH‡	<.103	3220	27	<.119	2053	17
IGFBP-3	<.107	1058	40	<.308	904	18
IGFBP-4	<.108	197	32	<.387	247	38
M-CSF	<.111	3	32	<.013	4	39
ANXA2‡	<.117	3131	37	<.070	1435	18
Osteopontin	<.155	17	17	<.019	45	40
IGFBP-1	<.158	1	14	<.098	3	30
IFN-γ	<.163	6	11	<.320	21	23
SCF	<.175	84	42	<.154	67	39
IGF-1	<.186	31	12	<.024	31	12
CA19-9	<.262	0	14	<.317	0	29
IGFBP-6	<.378	133	23	<.035	175	44

PDIA3, Protein disulfide isomerase A3; PER, peroxiredoxin; TIMP-1, tissue inhibitor of metalloproteinase 1; PHGDH, phosphoglycerate dehydrogenase; MTAP, methyl thioadenosine phosphorylase; IDH, isocitrate dehydrogenase; REC, recoverin; HSPA9, heat shock protein A9; ERP-29, endoplasmic reticulum protein 29; IP-10, interferon-γ-induced protein 10; GAPDH, glyceraldehyde 3 phosphate dehydrogenase; IGFBP-2, insulin-like growth factor-binding protein 2; SDF-1α, stromal cell-derived factor 1α; PGAM, phosphoglycerate mutase 1; HADH, hydroxacyl co-enzyme A dehydrogenase; ANXA1, annexin A1; IGFBP-7, insulin-like growth factor-binding protein 7; IMPDH, inosine 5'-monophosphate dehydrogenase; MYC, c-MYC; IL-2Rα, interleukin-2 receptor-α; IGFBP-5, insulin-like growth factor-binding protein 5; NYESO, NY-ESO; HSPA5, heat shock protein A5; CNN, calponin 2, isoform b; CYFRA 21-1, cytokeratin 19 fragment; TIM, triose isomerase mutase; FH, fumarate hydratase; IGFBP-3, insulin-like growth factor-binding protein 3; IGFBP-4, insulin-like growth factor-binding protein 4; M-CSF, monocyte colony-stimulating factor; ANXA2, annexin A2; IGFBP-1, insulin-like growth factor-binding protein 1; IFN-γ, interferon-γ; IGF-1, insulin-like growth factor 1; IGFBP-6, insulin-like growth factor-binding protein 6. *Log-rank. †Data presented as median fluorescence intensity for autoantibodies and as pg/mL for factors. ‡Autoantibody assay.

pathways represented are apoptosis and tumor suppression related (TIMP-1, survivin, and methylthioadenosine phosphorylase), metabolism by way of the glycolytic or

tricarboxylic acid cycles (α-enolase and isocitrate dehydrogenase), or protein synthesis and degradation (PHGDH, PDIA3, and ubiquilin). Because only 1 biomarker,

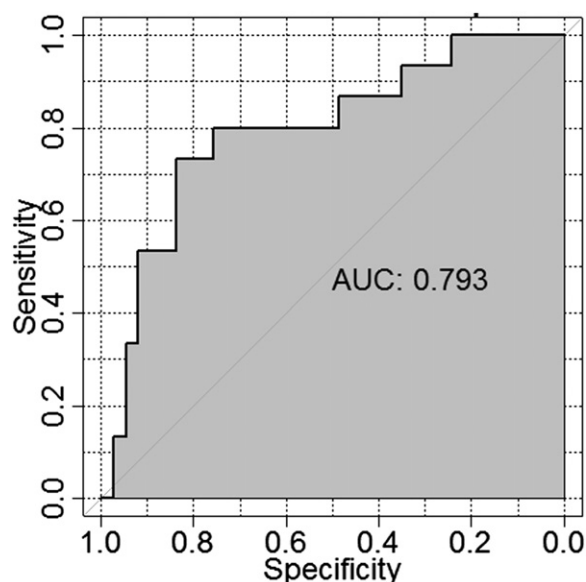


FIGURE 1. Receiver operating characteristic curve for 6-analyte algorithm for predicting recurrence in stage I NSCLC patients. Area under curve (AUC), 0.786, with sensitivity of 74% and specificity of 79%.

autoantibodies to ubiquitin, overlapped between the stage I groups and the lymph node-positive patients (Table 2), these markers were significant for recurrence alone and not grossly residual disease missed by surgery or pathologic examination. Hypothetically, patients selected using a simple and cost-effective serum test as having increased potential for recurrence would subsequently be candidates for neoadjuvant or adjuvant chemo- and/or radiotherapy and/or more frequent imaging during postoperative surveillance. When acted on early in the disease course, these additional interventions directed by the panel might shift the recurrence survival curve in a more favorable direction. Alternatively, this might also be a method to avoid unnecessary treatment in this same cohort.

PET-CT imaging is the current method of choice for the evaluation of metastatic progression before surgery and, for the purposes of predicting recurrence in stage I NSCLC, had performance characteristics of 74% sensitivity and 85% specificity in 1 study.¹⁰ Also, positive findings in 2 other studies showed that the maximum standard uptake value of 5.5 or more¹¹ or a standard uptake value of 4.7 or greater¹² was highly associated with recurrence. Although PET-CT has a similar performance to our biomarker panel, the vast difference in expense and exposure to radiation remain a concern for using PET-CT as a routine diagnostic test. The bulk of the studies in the arena of prognostic molecular classifiers for stage I NSCLC have thus far focused almost exclusively on the associations of gene expression microarray profiles from the primary tumor with pathologic lymph node status¹³⁻¹⁵ or clinical outcome¹⁶⁻¹⁸ (recently reviewed by Lin and Beer¹⁹). Examples of some of the

more distinguished efforts able to accurately stratify patients with good versus poor outcomes include studies by Raponi and colleagues,¹⁷ Shedden and colleagues,¹⁸ and Mitra and colleagues.²⁰ The most likely extension of these studies is translation of the expression signature into a quantitative reverse transcriptase-polymerase chain reaction-based assay for validation, with the intent of eventual implementation in clinical laboratories. Most recently, a 91-gene classifier algorithm was developed using 3 Affymetrix microarray data sets ($n = 680$) and translated into a quantitative real-time polymerase chain reaction assay for predicting patient survival. Validation of this 91-gene assay against 101 resected lung adenocarcinoma specimens resulted in an overall accuracy of 73.4%.²¹ However, despite the progress in this field, there has been some recent controversy in the published data asserting that tumor heterogeneity might lead to skewed (sampling-related) molecular profile information that, in turn, might provide misleading prognostic information.^{22,23} We reason that the evaluation of blood-based biomarkers for metastatic progression would be less prone to misclassifications of this type compared with tissue-based methods, given that the peripheral blood represents an integrative matrix of total tumor load and secondary tissues, with the latter possibly leading to potential signal amplification, such as in inflammation.

Our previous efforts in the development of multianalyte serum biomarker panels in NSCLC include several reports detailing the evolution of a serum test to accurately detect lymph node progression in surgical patients.^{6,8,9} Our last report presented a 6-analyte panel capable of classifying nodal status in resectable NSCLC patients with a sensitivity of 94%, a specificity of 97%, and an overall accuracy of 96%.⁸ As previously stated, for the present study, we reasoned that biomarkers capable of identifying patients with disease progression to the locoregional lymph nodes might also have utility in predicting recurrence in the stage I cohort. This hypothesis was based on the premise that occult micrometastases are the primary cause of disease recurrence in this cohort, implicating metastatic processes as a common mechanistic foundation to investigate. With this rationale, we also included patients with pathologically confirmed node-positive disease (pN1-N2) as “positive controls” for our studies. Although the performance of our “recurrence algorithm” (consisting of TIMP-1, IGFBP-4, and cytokeratin 19 fragment 21-1 and autoantibodies against α -enolase, heat shock protein A9, and PDIA3) ultimately had a cross-validated 77% accuracy for recurrence prediction and was not quite as robust as our “nodal algorithm,” we are very confident in our ability to further refine and improve this panel with our current efforts with shotgun proteomics. The data agree with our general hypothesis and show general trends of biomarker levels for the stage I cohort being more similar to those of the

lymph node-positive patients than the balance of stage I patients without recurrence. Finally, 1 of the more prominent studies on this topic was performed by D'Amico and colleagues²⁴ who evaluated 171 patients with NSCLC with 7 serum biomarkers (vascular endothelial growth factor, hepatocyte growth factor, E-selectin, CD44, basic fibroblast growth factor, urokinase plasminogen activator, and urokinase plasminogen activator receptor) using enzyme-linked immunosorbent assay. This was accomplished against pre- and postoperative (serial) blood sampling from patients undergoing complete resection. The Cox proportional hazards regression analysis showed disease recurrence could be predicted by decreasing levels of E-selectin ($P = .002$), increasing levels of CD44 ($P = .001$), and increasing levels of urokinase plasminogen activator receptor ($P = .03$), with serial sampling after surgery. A study is currently underway in our laboratory to determine whether the prognostic accuracy of our current panel can be improved with the findings by the D'Amico group and an effort to identify new biomarkers using proteomic methods. Furthermore, we will explore additional biomarkers for this purpose using both a shotgun proteomics approach and an approach focused on identifying surrogate biomarkers for the more prominent of the metastasis-suppressor genes, including *NM23*, *PEBP1*, and *KAI1*, as reviewed by Shoushtari and colleagues.²⁵

CONCLUSIONS

We have reported the development of a serum biomarker panel for stage I NSCLC patients that predicts recurrence previously undetected by standard imaging and pathologic studies. The performance characteristics of this initial "recurrence algorithm" (77% overall prognostic accuracy for 52 stage I NSCLC patients) were roughly in the same range as those for PET-CT and quantitative real-time polymerase chain reaction-based methods for predicting recurrence. Additional refinement of this panel has the potential to accurately stratify patients for more frequent screening protocols or adjuvant therapy. With validation of this panel against an external cohort, its true potential to improve overall survival can be realized.

References

1. American Cancer Society. Cancer facts and figures 2012. Available from: www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-031941.pdf. Accessed April 21, 2012.
2. Groome PA, Bolejack V, Crowley JJ, Kennedy C, Krasnik M, Sobin LH, et al. The IASLC Lung Cancer Staging Project: validation of the proposals for revision of the T, N, and M descriptors and consequent stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumours. *J Thorac Oncol*. 2007;2:694-705.
3. Flores RM, Ihekweazu UN, Rizk N, Dycoco J, Bains MS, Downey RJ, et al. Patterns of recurrence and incidence of second primary tumors after lobectomy by means of video-assisted thoracoscopic surgery (VATS) versus thoracotomy for lung cancer. *J Thorac Cardiovasc Surg*. 2009;141:59-64.
4. Lee BE, von Haag D, Lown T, Lau D, Calhoun R, Follette D. Advances in positron emission tomography technology have increased the need for surgical staging in non-small cell lung cancer. *J Thorac Cardiovasc Surg*. 2007;133:746-52.
5. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365:395-409.
6. Borgia JA, Basu S, Faber LP, Kim AW, Coon JS, Kaiser-Walters KA, et al. Establishment of a multi-analyte serum biomarker panel to identify lymph node metastases in non-small cell lung cancer. *J Thorac Oncol*. 2009;4:338-47.
7. Farlow EC, Vercillo MS, Coon JS, Basu S, Kim AW, Faber LP, et al. A multi-analyte serum test for the detection of non-small cell lung cancer. *Br J Cancer*. 2010;103:1221-8.
8. Patel K, Farlow EC, Kim AW, Lee BS, Basu S, Coon JS, et al. Enhancement of a multianalyte serum biomarker panel to identify lymph node metastases in non-small cell lung cancer with circulating autoantibody biomarkers. *Int J Cancer*. 2010;129:133-42.
9. Shersher DD, Vercillo MS, Fhied C, Basu S, Rouhi O, Mahon B, et al. Biomarkers of the insulin-like growth factor pathway predict progression and outcome in lung cancer. *Ann Thorac Surg*. 2011;92:1805-11.
10. Vansteenkiste J, Dooms C, De Leyn P. Early stage non-small-cell lung cancer: challenges in staging and adjuvant treatment: evidence-based staging. *Ann Oncol*. 2010;21(Suppl. 7):vii189-95.
11. Goodgame B, Pillot GA, Yang Z, Shriki J, Meyers BF, Zoole J, et al. Prognostic value of preoperative positron emission tomography in resected stage I non-small cell lung cancer. *J Thorac Oncol*. 2008;3:130-4.
12. Shiono S, Abiko M, Sato T. Positron emission tomography/computed tomography and lymphovascular invasion predict recurrence in stage I lung cancers. *J Thorac Oncol*. 2011;6:43-7.
13. Seike M, Yanaihara N, Bowman ED, Zanetti KA, Budhu A, Kumamoto K, et al. Use of a cytokine gene expression signature in lung adenocarcinoma and the surrounding tissue as a prognostic classifier. *J Natl Cancer Inst*. 2007;99:1257-69.
14. Takada M, Tada M, Tamoto E, Kawakami A, Murakawa K, Shindoh G, et al. Prediction of lymph node metastasis by analysis of gene expression profiles in non-small cell lung cancer. *J Surg Res*. 2004;122:61-9.
15. Xi L, Coello MC, Litle VR, Raja S, Gooding WE, Yousem SA, et al. A combination of molecular markers accurately detects lymph node metastasis in non-small cell lung cancer patients. *Clin Cancer Res*. 2006;12:2484-91.
16. Beer DG, Kardias SL, Huang CC, Giordano TJ, Levin AM, Misek DE, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med*. 2002;8:816-24.
17. Raponi M, Zhang Y, Yu J, Chen G, Lee G, Taylor JM, et al. Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res*. 2006;66:7466-72.
18. Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, Gerald WL, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med*. 2008;14:822-7.
19. Lin J, Beer DG. Molecular predictors of prognosis in lung cancer. *Ann Surg Oncol*. 2012;19:669-76.
20. Mitra R, Lee J, Jo J, Milani M, McClintick JN, Edenberg HJ, et al. Prediction of postoperative recurrence-free survival in non-small cell lung cancer by using an internationally validated gene expression model. *Clin Cancer Res*. 2011;17:2934-46.
21. Chen G, Kim S, Taylor JM, Wang Z, Lee O, Ramnath N, et al. Development and validation of a quantitative real-time polymerase chain reaction classifier for lung cancer prognosis. *J Thorac Oncol*. 2011;6:1481-7.
22. Reck M, Hermes A, Tan EH, Felip E, Klughammer B, Baselga J. Tissue sampling in lung cancer: a review in light of the MERIT experience. *Lung Cancer*. 2011;74:1-6.
23. Thunnissen E, Kerr KM, Herth FJ, Lantuejoul S, Papotti M, Rintoul RC, et al. The challenge of NSCLC diagnosis and predictive analysis on small samples: practical approach of a working group. *Lung Cancer*. 2011;76:1-18.
24. D'Amico TA, Brooks KR, Joshi MB, Conlon D, Herndon J II, Petersen RP, et al. Serum protein expression predicts recurrence in patients with early-stage lung cancer after resection. *Ann Thorac Surg*. 2006;81:1982-7.
25. Shoushtari AN, Szmulewitz RZ, Rinker-Schaeffer CW. Metastasis-suppressor genes in clinical practice: lost in translation? *Nat Rev Clin Oncol*. 2011;8:333-42.

Discussion

Dr Chuong D. Hoang (Stanford, Calif). I want to congratulate you and your group on a provocative research effort, and, if your ongoing work demonstrates that a blood-based test can accurately predict disease recurrence in early-stage lung cancer, this could be

an incredible opportunity to identify novel mechanistic insights to lung cancer development and progression.

I have the following 2 questions based on your presentation. First, the reasoning for your studies is that occult micrometastases undetectable at surgery are the primary cause of disease recurrence in early-stage lung cancer patients. With this in mind, why would you want to collect blood samples for a predictive biomarker assay in the preoperative period? The results of the blood assays could be confounded by the presence of the primary tumor, which might obscure the signal of the micrometastases. Why not collect the blood samples in the postoperative period when the effects of the primary tumor are no longer present?

Second, the primary method you used to select your predictive analyses and the analytes was the Random Forest algorithm. This classifier tends to work best when the number of input variables is large, and yet you only tested 43 biomarkers for this study. What were the selection criteria you used for including these potential biomarkers? Also, this algorithm suffers sometimes from overfitting of the data, which can lead to near-perfect classifier results, and I noticed that you showed relatively lower sensitivity, specificity, and accuracy in your final analysis.

Again, I really look forward to seeing your validation studies on this interesting project. I thank you for the opportunity to discuss it.

Dr Rinewalt. Thank you, Dr Hoang, for your input and for your questions.

To answer your first question, we chose to do these preoperatively because this was based on our previous discovery efforts. We know that these tumor antigens typically fall into 3 categories: aberrant isoforms of commonly expressed proteins, overexpression of common proteins, or proteins specific to the tumor itself, the latter of which is what most of us are specifically interested in targeting for either diagnosis or treatment. These are typically only present in minuscule quantities, even from the primary tumor. We hypothesized that we would obtain a greater degree of sensitivity by doing assays at this time for the circulating autoantibodies to the tumor instead of the tumor antigens directly, given the large signal implication produced by the B cells of the body's own humoral response. Additionally, autoantibodies are thought to only be produced early on in the disease process.

We selected these biomarkers for the study using 2 methods. The first was based on reviews of the published data for biomarkers that had previously demonstrated some efficacy for the early detection of lung cancer. The second was based on our previous discovery efforts in our laboratory, headed by Drs Michael Liptay and Jeffrey Borgia. Specifically, we used Western blots prepared from

a lung adenocarcinoma cell line and then probed these with pooled serum from non-small cell lung cancer patients, 1 without nodal disease and 1 with nodal disease. We compared these and identified them using mass spectrometry to choose this list of biomarkers.

To answer your questions about the Random Forest, it is a very complicated multivariate statistical method. You are correct in that it has been shown in the published data to be a robust and highly accurate when one is analyzing data with a large number of inputs, and we did have 43 biomarkers. Less noise is present when analyzing this type of data. We believe that this will have resulted in less overfitting of the data than, say, using just a classification tree, because the Random Forest will use approximately 1000 trees and then average their results.

Dr Raphael Bueno (*Boston, Mass*). I congratulate you on your paper.

I have 2 questions. One, you are looking for survivors versus patients likely to develop recurrence so you can affect treatment, so I would expect your curves, the split curve, to look as a curve would with survival at 5 years of 95% and a curve with survival of 20% at 5 years, which is really what you need achieve. Any comments on that?

Second, we do not live in isolation. The pathologists have defined, and we have defined, during the past 3 decades, as surgeons, prognostic factors from studying the tumor, from the differentiation, from the T status, for lymphovascular invasion, et cetera, that are well established, and any prognostic test needs to be put, as a part of that, in the context of that to bring an additional piece of information. You do not merely want to recapitulate T1 versus T2. Have you considered both issues?

Dr Rinewalt. Thank you for your comments, Dr Bueno.

To answer your first question, I agree that ultimately the goal is to try to prognosticate these patients, and so we would hope that at some point we will be able to develop a curve with which we can predict who will develop recurrence and hopefully stratify them into a new treatment that will shift that first curve up. The patients who present with distant disease will probably continue to do poorly; however, this was not the goal of our study. We are hoping that by being able to intervene early on in the process that we can shift that survival curve up.

To answer your second question, we have not specifically put these together with all of the clinical factors, although I think ultimately studying all the data that we have, including PET data, tumor size, and histologic type, this will be a companion test and not something that will stand alone.